ARI Research Note 91-58

AD-A238 941

The Relationship of Simulator Fidelity to Task and Performance Variables

John Allen, Louis Buffardi, and Robert Hays
George Mason University

for

Contracting Officer's Representative
Michael Drillings

Office of Basic Research
Michael Kaplan, Director

June 1991

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Research accomplished under contract
for the Department of the Army

Department of Psychology, George Mason University

Technical review by

Michael Drillings

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS
--

2a. SECURITY CLASSIFICATION AUTHORITY
--

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE
--

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
--

5. MONITORING ORGANIZATION REPORT NUMBER(S)
ARI Research Note 91-58

6a. NAME OF PERFORMING ORGANIZATION
Department of Psychology
George Mason University

6b. OFFICE SYMBOL *(If applicable)*
--

7a. NAME OF MONITORING ORGANIZATION
U.S. Army Research Institute
Office of Basic Research

6c. ADDRESS *(City, State, and ZIP Code)*
Fairfax, VA 22030-4444

7b. ADDRESS *(City, State, and ZIP Code)*
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
U.S. Army Research Institute for the Behavioral and Social Sciences

8b. OFFICE SYMBOL *(If applicable)*
PERI-BR

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
MDA903-82-K-0464

8c. ADDRESS *(City, State, and ZIP Code)*
Office of Basic Research
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102B	74F	N/A	N/A

11. TITLE *(Include Security Classification)*
The Relationship of Simulator Fidelity to Task and Performance Variables

12. PERSONAL AUTHOR(S)
Allen, John; Buffardi, Louis; and Hays, Robert

13a. TYPE OF REPORT
Final

13b. TIME COVERED
FROM 82/06 TO 85/06

14. DATE OF REPORT *(Year, Month, Day)*
1991, June

15. PAGE COUNT
96

16. SUPPLEMENTARY NOTATION
Michael Drillings, Contracting Officer's Representative

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
~~Simulator fidelity~~ Transfer of training
~~Training devices~~ Fidelity specifications
*Maintenance training (Continued)

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This report describes a 3-year program of basic research to investigate the relationship between simulator fidelity and maintenance training effectiveness. A unique feature of this work was the use of a specially-designed reference system and simulator testbed for use in training simple electromechanical troubleshooting skills. Researchers report two experiments designed to compare the transfer performance of subjects in nine different simulator training conditions. In both studies two aspects of simulator fidelity were manipulated--the degree to which a training simulator "looked like" actual equipment (physical fidelity), and the degree to which it "acted like" real equipment (functional fidelity). Results led to the following conclusions.

*Simulators, *Training devices

1. Simulator fidelity should not be considered a single, uniform concept, but a multi-dimensional one consisting of at least a physical and functional component.

(Continued)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Michael Drillings

22b. TELEPHONE *(Include Area Code)*
(703) 274-8722

22c. OFFICE SYMBOL
PERI-BR

DD Form 1473, JUN 86 *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

ARI Research Note 91-58

18. SUBJECT TERMS (Continued)

Fidelity requirements,
Troubleshooting,
Simulators.

19. ABSTRACT (Continued)

2. With "cognitive" tasks like troubleshooting, the functional similarity of training simulators to real equipment is a critical variable.

3. Temporal variables such as time-to-solution, intertest time, and time-to-first attempted solution are particularly sensitive to fidelity manipulations.

4. In simple troubleshooting tasks that emphasize cognitive skills, there may not be an advantage to "actual equipment" training.

5. The difficulty level of criterion problems during transfer affects performance in a general way independent of simulator fidelity levels experienced during training.

6. A number of individual difference variables seem to correlate well with the transfer of troubleshooting performance, e.g., GRE-A scores and tests of intellectual and social interests.

7. Interactions between individual difference variables and training conditions suggest that the concept of adaptive training may be an applicable one in training situations like the one examined.

Researchers discuss recommendations and suggestions for future research.

ACKNOWLEDGMENTS

The authors wish to thank the following George Mason University faculty members and students for their valuable contributions to the work described in this project: Dr. Evans Mandes, Dr. Bob Pāsnak, Dr. Jim Sanford, Dr. Zita Tyer, Mark Aceto, Alex Ayres, Rack Campbell, Annette Cross, Beth Drum, Rob Huggins, Rick Hall, Joe Maroun, Lou Matzel, and Ray Moffett. Special thanks are also due Margaret Baechtold not only for her special care in helping put together the various reports required by the project, but also for her cheerful and untiring help in managing the day-to-day details of the project.

THE RELATIONSHIP OF SIMULATOR FIDELITY TO TASK AND PERFORMANCE VARIABLES

EXECUTIVE SUMMARY

Requirement:

To undertake a 3-year program of basic research to investigate the degree of functional and physical fidelity needed in maintenance training simulators to ensure maximum transfer of training. To accomplish this by constructing and using a reference system and simulator testbed for simulation and actual equipment training. Also, to initiate studies designed to examine the interactive role of such variables as trainee aptitudes and abilities, problem complexity, and other training systems variables.

Procedure:

A special simulation and training testbed for use in training simple electromechanical troubleshooting skills was designed and constructed and used to investigate two aspects of training device fidelity--the degree to which a training simulator "looked like" actual equipment (physical fidelity), and the degree to which it "acted like" real equipment (functional fidelity). When baseline data and pilot work was completed on the reference system and simulators, two experiments designed to compare the transfer performance of subjects in nine different simulator training conditions were conducted. Transfer performance was assessed in terms of a number of dependent measures, and transfer scores were correlated with a variety of individual difference variables.

Findings:

First, the results indicated that simulator fidelity should not be considered a one-dimensional concept but a multidimensional one consisting of at least a physical and functional component. Second, the fact that differential performance effects sometimes occurred as a function of the dependent variable selected points to the importance of choosing adequate criterion performance measures for assessing transfer. Third, with "cognitive" tasks such as the one used, functional similarity between the training simulator and its reference system is much more important than its physical similarity. Fourth, temporal variables such as time between actions, time-to-solution,

and time-to-first solution are especially sensitive to fidelity manipulations, and this is particularly true with respect to functional fidelity. Fifth, in simple troubleshooting tasks that emphasize learning cognitive skills, there may not be any advantage to actual equipment training. Sixth, the fact that a number of significant interactions were noted between certain individual difference variables and training conditions strongly suggests that the concept of adaptive training may be an especially useful one in training situations like the present one.

Utilization of Findings:

In addition to entry into the Army's computerized database on fidelity and other training systems issues, the present findings may be of immediate value to both military and civilian training systems designers involved in the design, development, and implementation of simulator and other training devices. Aside from the obvious relevance to maintenance training, teaching of troubleshooting skills, and design of maintenance simulators, these findings may also be of significant value for future research in the area.

THE RELATIONSHIP OF SIMULATOR FIDELITY TO TASK AND PERFORMANCE VARIABLES

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

THE RELATIONSHIP OF SIMULATOR FIDELITY
TO TASK AND PERFORMANCE VARIABLES

I. PURPOSE OF THE PRESENT REPORT

The aim of the present report is to describe and summarize the accomplishments of the STAR (Simulation, Training and Research) Laboratory, Department of Psychology, George Mason University, on the U.S. Army Research Institute Contract MDA 903-82-K-0464 ("The Relationship of Simulator Fidelity to Task and Performance Variables"). Work began on this project on June 7, 1982, and was completed on June 7, 1985.

II. BRIEF DESCRIPTION OF PROJECT

Although a full account of the present contract and its objectives can be found in the original proposal, it may be briefly summarized as follows.

The current work is part of a program of basic research aimed at investigating the degree of functional and physical fidelity needed in maintenance simulators to ensure maximum transfer of training. A unique feature of this research is the construction and use of a generic reference system (Actual Equipment Trainer, i.e., AET) for simulation and actual equipment training. A major objective is to provide data upon which fidelity-value guidelines for maintenance training devices and simulators might be based. Studies aimed at answering questions about fidelity levels and involving training in a relatively simple electro-mechanical troubleshooting task were conducted during the period.

Major features of the project include:

--Development of operational definitions of physical and functional fidelity

--Design and construction of a reference system for simulation and actual equipment training (AET)

--Collection of baseline troubleshooting data on the reference operating system (AET)

--Comparison of troubleshooting training performance on several simulator variations (with differing degrees of functional and physical fidelity) with that of the reference system

--Parametric investigation of simulator training effectiveness using transfer of training paradigms.

III. Background

A. Simulator Fidelity

The issue of simulator fidelity has been discussed and studied for over thirty years now. During this time, the term has been used in a variety of ways and to refer to many different aspects of the training situation (Hays, 1980). A representative sample includes such things as: equipment fidelity, environmental fidelity, psychological fidelity, task fidelity, physical fidelity, functional fidelity, similarity, degree of correspondence, and so on. A common thread, however, is that together they imply at least two major features or dimensions along which training simulators may differ from actual equipment, i.e., in terms of how they "appear" (physical fidelity) and in terms of what they "do" (functional fidelity). Because these two features seem to capture most of the spirit of much of what is meant by fidelity, we have chosen to focus on them in our work.

A major problem, of course, is that of defining what is meant by the concepts of physical and functional fidelity. If,

however, one accepts Hays' (1980) definition, then straightforward operations can be delineated. That is, physical fidelity indexes the degree to which a simulation physically reproduces the appearance and physical activities of a system. From this perspective, there are at least three potential levels of physical fidelity. High physical fidelity consists of a simulation which incorporates the complete reference system. Such simulation includes all of the physical complexities of the simulated system, although the environment itself may be different or change. This type of simulation, while maximally familiarizing the trainee with the "look" and "feel" of the system, is relatively expensive. Also, the complexities may delay learning, and expensive components may be fragile and easily destroyed.

Moderate physical fidelity consists of a modified version of the reference system, which does not provide all of the physical "gestalt" or appearance of the system. Moderate physical fidelity is provided, for example, by a portion of the system. Some components may be removed, or their physical appearance altered. A side benefit of this simplification is that it may aid learning by isolating key elements to which a trainee must attend.

Low physical fidelity represents the final level of simplification. The physical appearance of the simulator may bear no resemblance to the reference system. Low physical fidelity may be represented, for example, by a two-dimensional schematic or functional diagram. While a paper-and-ink drawing may only symbolize the reference system, the sequence and inter-dependence of component parts may be faithfully represented. This may allow a trainee ample opportunity to learn problem-

solving strategies. This, in turn, can be a productive first step toward quick and efficient learning.

High functional fidelity would consist of having all the output components/feedback of the reference system operating. For moderate functional fidelity, certain components of the system would be capable of operating and providing feedback to the trainee; however, others would not. Finally, with low functional fidelity, the system components would simply not work or provide feedback at all. These relationships are more clearly depicted in Table 1.

B. <u>Fidelity, Transfer of Training and Maintenance Training Effectiveness</u>

As Hays (1980) has pointed out, the problem of determining proper levels of simulation does not stand in isolation, and its relationship to training effectiveness (i.e., how well training on a simulator transfer to work on operational equipment) is not a simple one. In fact, there are a host of variables and factors which must be considered, i.e., the context within which the trainer is used, the kind of task trained, the stages of learning involved, learner abilities and capabilities, task difficulty, the effects of various instructional features, etc., to name a few. While Adams (1979) has pointed out some difficulties with such a design, the fact that it permits one to easily gauge the relative transfer of learning from one system to another under well-controlled experimental conditions make it a logical choice. Too, as Hays (1980) has pointed out, it permits the derivation of a variety of measures to assess training effectiveness, such as the transfer effectiveness ratio used by Provenmire and Roscoe

TABLE 1

Physical and Functional Fidelity Relationships and Simulation Conditions (GMU Test Bed)

		Physical Fidelity High	Physical Fidelity Medium	Physical Fidelity Low
Functional Fidelity	High	All relays and output devices present Hand-held unit provides feedback; output devices function	Some relays and outputs real; some mockups Hand-held unit provides feedback; output devices or symbols function	Pen-and-Ink drawing Hand-held provides feedback; symbols function
Functional Fidelity	Medium	All relays and output devices present Hand-held unit provides feedback; output devices do not function	Some relays outputs real; some mockups Hand-held unit provides feedback; output devices do not function	Pen-and-Ink drawing Hand-held unit provides feedback; outputs do not work
Functional Fidelity	Low	All relays and output devices present Hand-held unit does not provide feedback; output devices do not function	Some relays and outputs real; some mockups Hand-held unit does not provide feedback; output devices do not function	Pen-and-Ink drawing Hand-held unit does not provide feedback; output devices do not work

(1971). (Note: Various transfer designs and other measures of transfer, along with a discussion of their advantages and disadvantages, have been neatly summarized by Ellis (1965).)

Lest it be lost among discussions of transfer measures, designs, fidelity levels, etc., it should be pointed out that the crucial question to be addressed is how much transfer of training is associated with training on simulators which vary in their resemblance to the reference operating system. The important question, as others have pointed out, is how to achieve training, not realism, (Bunker, 1978). Thus, if optimizing training is made the real objective, specification of proper levels of simulator fidelity becomes partially determined by the amount of transfer desired. (Kinkade & Wheaton, 1972).

C. Troubleshooting Tasks

One of the most ubiquitous activities performed by maintenance personnel is troubleshooting or fault diagnosis. Indeed, the kind of problem-solving skills behavior involved in locating a malfunction in a turbo-prop engine, an automatic pilot, a piece of electronic communication gear, etc., are similar. In short, a maintenance person must identify a suspected component and interpret its status relative to knowledge of overall system operation, system outputs, previous component checks and readings, etc., and come to a decision as to what component(s) are faulty or malfunctioning. Because of the prevalence and importance of this activity, our research has focused on training the troubleshooting skills involved in diagnosing problems in an electro-mechanical/hydraulic operating system.

As for troubleshooting per se, it may be viewed as a special kind of problem-solving in which particular symptoms of a device demand that certain decisions be made by the maintenance person which lead to correct identification of the system malfunction. In terms of skills, troubleshooting relies primarily, if not almost exclusively, on cognitive processes. Such skills may be contrasted with perceptual-motor skills which involve combinations of sensory, perceptual, and manipulative components. It hardly need be said that almost all military maintenance tasks involve a blend of both categories of skill, however, for purposes of analysis it is helpful to consider them separately. That is, not only must maintenance personnel typically identify and interpret symptoms (cognitive skills), but he/she must also adjust screws, align guages, remove and replace suspect components, etc. (perceptual-motor skills). As far as knowledge is concerned, however, there seems to be far less known about the relationship(s) between device fidelity and cognitive maintenance skills than for perceptual-motor skills and fidelity (Baum et al, 1982). Such a state of affairs points to a continuing need for studies which are systematically aimed at filling in this knowledge gap. Until we know more, however, the amount of training simulator fidelity needed to assure maximum transfer of training will be subject to "guesswork."

We believe a major advantage of the reported research is that it focuses on training the cognitive side of the maintenance task in such a way as to "shed light" on the nature of the simulator fidelity-troubleshooting function. The fact that the reference

operating system used (which represents the actual equipment to-be-simulated) is easy to fault lends itself nicely to the collection of transfer of training data; also, comparisons between simulator performance and actual equipment are very straightforward. This has not always been the case in previous work (Cicchivelli, 1980).

IV. SUMMARY OF MAJOR ACCOMPLISHMENTS DURING THE FIRST YEAR

Major accomplishments during the first year of the contract may be summarized as follows:

--Design, construction, testing and evaluation of a generic "reference system" (comprised of electro-mechanical/hydraulic components and various output devices) for use as an actual equipment trainer (AET) and for simulation/modeling

--Design, construction, testing, and evaluation of reference system support equipment and interface devices

--Collection of preliminary baseline data on the reference system and low functional/low physical simulator

--Development and evaluation of preliminary instructions to troubleshooting trainees

--Design and construction of low functional/low physical, low functional/high physical and medium functional/medium physical simulators and associated equipment

--Preliminary design of high functional/low physical simulator

--Completion of behavioral and task analyses of trainee performance on the reference system

--Preliminary evaluation of baseline data on reference system

--Development, testing and evaluation of data-reduction software for the reference system

--Development and testing of an Apple computer software program for use in the collecting and timing of observations during simulator training and behavioral/task analyses

--Preliminary identification of training and "test probe" problems for use in initial pilot transfer studies

--Establishment and organization of reference library of materials relating to simulator fidelity, maintenance training, fault diagnosis, troubleshooting and training devices in general.

A more complete discussion of these activities and accomplishments may be found in the annual report for the first year of the project.

V. SUMMARY OF MAJOR ACCOMPLISHMENTS AND PROGRESS DURING YEAR 2

Major accomplishments and progress during the second year of the project may be summarized as follows:

--Substantial refinement of the reference system and its associated equipment was completed

--The medium and low physical fidelity simulators were completed and tested

--Baseline data collection on the reference system was continued

--Considerable refinement of the basic 'SIMMYD' data acquisition and recording software was completed

--Identification of training and transfer problems for use during transfer studies was completed

--Instructions to be given to subjects during transfer studies were developed

--A battery of aptitude, attitude, and ability tests was identified and selected for use during the transfer experiments

--A pilot study focusing on the transfer between simulator training and problem solving on the reference system was designed and completed in the Fall, 1983

--The first major transfer study involving nine training simulator conditions and the reference system was begun during the Spring, 1984

--A 'simulator affect scale' was developed for use as a tool to help scale the functional and physical fidelity dimensions.

A more detailed discussion of some of these activities and accomplishments may be found in the annual report for the second year of the project.

VI. Objectives and Goals for the Third Year of the Project

Two studies, each of which compared the relative training effectiveness of nine simulator-based training conditions differing in terms of their physical and functional simularity to the reference system, were planned. A detailed description of these studies and a discussion of their outcomes is presented in the present report.

VII. Major Accomplishments During the Third Year

A. Transfer Study I

A.1 Background Purpose and Rationale of Experiment I

Previous research has, for the most part, been concerned with evaluating the training effectiveness of full-fidelity devices rather than systematically investigating the effect of various degrees of similarity (Ayres, Hays, Singer & Heinicke,

1984). Notable exceptions are the research efforts of Wheaton and Mirabella (1972), Mirabella and Wheaton (1974), Johnson and Rouse (1982), Baum, Reidel, Hays and Mirabella (1982) and Johnson and Fath (1983). Each study systematically manipulated the characteristics of the devices used to train the criterion tasks. Of these, however, only Baum et al. (1982) varied both the physical and functional characteristics of the training simulators.

In the Baum et al, (1982) study, two aspects of fidelity were manipulated, viz., the degree to which a training simulator "looked like" the actual equipment it was simulating (physical fidelity), and the degree to which it "acted like" real equipment (functional fidelity). Training simulators of various degrees of physical and functional fidelity were used to train a simple mechanical adjustment task. Although they noted a significant effect for physical fidelity, no main effect for functional fidelity or interaction effects between the two fidelity dimensions were found. Based on these results, they concluded that, at least for simple adjustment tasks, high physical similarity may be very important from the standpoint of maximum learning transfer. They proposed additional research using different tasks to further specify the effects of physical and functional fidelity. In part, the present study may be viewed as a response to this suggestion since one of the major aims was to investigate the effects of simulator fidelity during the training of a relatively simple troubleshooting task. An additional aim was to collect data on a number of individual difference variables to help determine how such variables might interact with fidelity during training.

Literature reporting individual difference predictors of troubleshooting performance is quite sparse. Henneman & Rouse (1984), taking advantage of test information already available on subjects participating in a troubleshooting experiment, investigated the relationships of general academic ability (ACT scores, GPA), mechanical aptitude (Survey of Mechanical Insight), and cognitive style (reflective-impulsive, field dependent-independent) with performance. They found evidence that cognitive style and academic ability significantly correlate with troubleshooting performance. Rouse & Rouse (1982) confirmed the finding for cognitive style, in that reflective and field independent individuals were more effective at fault diagnosis.

The present study took a broader and more systematic approach than previous research by investigating a wide array of individual differences that might be related to not only troubleshooting performance, but also to different modes of training people on such tasks. Thus, measures were chosen to reflect the logical, analytic abilities and general interests apparently required by the task.

The present research used a specially-constructed test bed device for high fidelity simulation training and transfer of training testing. (See Year 2 annual report for a more detailed description of the test bed.) Degraded simulations of this device produced the various levels of physical and functional fidelity used to investigate the effects of fidelity on transfer of training. Physical fidelity manipulations involved variations in the way components and their spacial relationships were represented in the simulator. Functional fidelity, on the other hand, was

defined in terms of the degree of informational feedback (Bilodeau, 1966) available to the subject, or what might be called the informational aspect of equipment function. This aspect of functional fidelity may be contrasted with its "stimulus/response options" aspect which refers to opportunities the equipment provides the subject to receive stimuli (e.g., a dial moves) and/or give responses (e.g., the subject can turn a knob) (Hays, 1980).

A wide variety of performance measures were chosen to investigate fidelity effects. This strategy was chosen both because of the exploratory nature of the research and also because previous studies have shown that no one variable fully describes how subjects solve troubleshooting problems (Glass, 1967; Finch, 1971).

A.2 Method

A.2.1 Subjects

One hundred college undergraduates (40 males and 60 females) ranging in age from 17 to 55 were drawn primarily from introductory psychology courses at George Mason University and served as paid subjects ($5.00/hour). Subjects also received course credit for participation.

A.2.2 Materials

Individual differences measures. All subjects completed a two-hour test battery which included the following measures: (1) Group Embedded Figures Test (GEFT), (2) Bennett Mechanical Comprehension Test (BMC), (3) Graduate Record Examination - Analytic (GRE-A), (4) Rotter's Locus of Control Scale (LOC) and (5) Holland's Vocational Preference Inventory (VPI). The VPI

consists of 11 scales: (1) Realistic (VPI-R), (2) Intellectual (VPI-I), (3) Social (VPI-S), (4) Conventional (VPI-C), (5) Enterprising (VPI-E), (6) Artistic (VPI-A), (7) Self-Control (VPI-SC), (8) Masculinity (VPI-M), (9) Status (VPI-ST), (10) Infrequency (VPI-F), and (11) Acquiscence (VPI-AC).

MINC-11 Computer. All experimental activities were controlled and monitored with a MINC-11/23 (Digital Equipment Corp.) computer. The computer controlled the reference system and associated simulators and also recorded and stored all subject actions during the training and testing phases of the experiment.

Reference System. The reference system, which served as both the "actual equipment" and as the high physical fidelity trainer, may be described as follows. Twenty-eight electromechanical relays and five solid-state pullup panels were interconnected to eight output devices (viz., fan, water pump, solenoid valve assembly, three lights, TV monitor and sound generator and speaker). When operating properly (i.e., when no faults had been introduced into the system) all of the output devices worked. However, when a relay or pullup panel was faulted by the experimenter, associated relays/pullup panels and one or more of the output devices did not work. A subject's job was to discover which relay or pullup panel assembly was at fault. To do this, the subject had to first detect which output device(s) were not working, and then, by testing individual relays and pullup panels, discover which component was at fault.

Hand-held Tester. Relays were tested by means of a specially - designed hand-held tester. To test a relay with this device, the trainee dialed in the number of the relay, placed the probe

on the relay checkpoint, and then pressed a test button located on top of the unit. If the component under consideration was operating properly, a green light illuminated on the face of the tester. If the relay was not working, the green light did not light. A cable connecting the hand-held tester to the MINC-11 computer was used to signal when a test had been requested. Recording of the number and types of responses and the times between responses was done automatically.

Faulter Panels. Individual relays and pullup panels were faulted by means of a two faulter panels (one for the reference system and one for the medium and low physical fidelity simulators) located in the experimenter's room. These faulter panels consisted of a number of lights and switches, each of which corresponded to a particular relay or pullup panel assembly. To fault a component, the experimenter merely threw the switch appropriate to that component. This action caused the component to stop working. All interconnected relays, pullup panels and output devices then ceased to function. Lamps on the face of the panels provided the experimenter instant information as to which component had been faulted.

Simulators. During the training phase, three major training simulators, each differing along the physical fidelity dimension, were used (i.e., high, medium and low). These simulators, when coupled with one of three levels of functional fidelity (i.e., high, medium or low), provided nine possible training simulators.

High Physical Fidelity Simulator. This simulator was the reference system, as described above.

Medium Physical Fidelity Simulator. The medium fidelity simulator was quite similar to the reference system in terms of its size and general appearance. Half of its components (relays, output devices, etc.) actually duplicated those found on the reference system whereas half were wooden mockups. Thus, trainees using this system were able to test and/or observe the responses of some components and devices, but not others.

Low Physical Fidelity Simulator. Visually, the low physical fidelity simulator was different from either the medium or high physical fidelity simulators. Essentially, it was a symbolic representation of the reference system. Relays, pullup panels and output devices were depicted by rectangles, and lines between rectangles indicated wiring connections. Testing and checking of relays and output devices (if required in a particular training condition) could be accomplished by using the hand-held tester.

A.2.3 Design and Procedure.

During the first session subjects were tested in groups, of five to ten, on the battery of individual difference measures listed above. Testing lasted approximately two hours.

Approximately one week after the testing phase, subjects were assigned to one of ten training groups shown in Table 2. Subjects were assigned to groups in random order until 10 subjects were trained in each group.

At the beginning of training, a subject listened to one of nine versions of taped instructions which explained the use of the apparatus and the general intent of the training. Subjects

TABLE 2

Experimental Design: Experiments I and II

		Physical Fidelity		
		High	Medium	Low
Functional Fidelity	High	All relays and output devices present Hand-held unit provides feedback: output devices function	Some relays and outputs real; some mockups Hand-held unit provides feedback; output devices or symbols function	Schematic drawing Hand-held provides feedback; symbols function
	Medium	All relays and output devices present Hand-held unit provides feedback; output devices do not function	Some relays outputs real; some mockups Hand-held unit provides feedback; output devices do not function	Schematic drawing Hand-held unit provides feedback; outputs do not work
	Low	All relays and output devices present Hand-held unit does not provide feedback; output devices do not function	Some relays and outputs real; some mockups Hand-held unit does not provide feedback; output devices do not function	Schematic drawing Hand-held unit does not provide feedback; output devices do not work

No-Training Control Group

in the no-training control group also heard one of the nine instructions sets (assigned randomly), and then read from a set of non-related psychology readings for a period of time equal to that taken by subjects in the simulator training conditions. During the taped instructions, the experimenter remained in the room to demonstrate the various equipment operations and to answer any questions. When instructions were completed, the experimenter left the room.

During training, subjects received eight training problems (or "faults). Three sets of eight problems each were used. Sets were matched in terms of problem difficulty based on pilot data. Each subject was trained on one randomly determined set.

When a training problem was begun (through the initiation of a fault by the experimenter), the subject was signalled by a start tone. The subject's task was then to determine which of the possible 28 components was at fault. Depending on the simulator in use, this either involved testing relays and output devices (high functional fidelity), testing relays only (medium functional fidelity) or, as was the case in the low functional fidelity conditions, merely examining and studying components and their relationships. An interval of approximately 30 seconds occurred between each problem. After training, the subject received a 5 minute break.

After the break, the transfer portion of the session was begun. First, however, subjects were introduced to the reference system via a second set of recorded instructions which were identical to those heard by subjects in the high physical/high functional training condition.

During the transfer test, subjects were asked to solve six new problems on the reference system. All subjects solved the same set of six problems, with the presentation order randomized. Degree of transfer was assessed using the following dependent variables: time-to-first solution (TFS), inter-response time (IRT), which refers to the time between any two actions either tests or solutions, inter-test time (ITT), time-to-correct solution (TS), number of tests (T), number of tests repeated (TR), number of attempted solutions (S), and number of solutions repeated (RS).

The training and transfer phase lasted approximately one and one half to two hours.

A.3 Results and Discussion

A.3.1 ANOVAs on Dependent Variables

Three x three factorial analyses of variance comparing the transfer performances of the groups trained in each fidelity condition were performed on each of the eight dependent variables. Means and standard deviations for the various groups are shown in Table 3.

A significant main effect for both physical fidelity $F(2,81)=3.24$, $p<.04$ and functional fidelity, $F(2,81)=9.89$, $p<.0001$ was found for the total-time-to-solution variable. (See Table 4). Persons trained on devices with high physical and high functional fidelity reached correct solutions in less time than persons trained in lower fidelity conditions. The main effect of functional fidelity manipulations was found to be significant for time-to-first-solution $F(2,81)=3.15$, $p<.05$; inter-test time $F(2,81)=8.44$, $p<.001$; inter-response time $F(2,81)=9.68$, $p<.0002$; and number of repeated tests $F(2,81)=5.48$, $p<.01$. Thus, for all of these measures, higher

TABLE 3

Means and Standard Deviations for Dependent Variables on Transfer Test

Variables		Training Condition (physical fidelity/functional fidelity)								
		HH	MH	LH	HM	MM	LM	HL	ML	LL
Number of Tests	M	7.93	8.06	6.73	6.38	9.65	10.62	10.06	8.71	6.88
	SD	5.10	3.17	2.37	2.94	5.57	4.74	4.15	6.33	3.14
Number of Solutions	M	3.16	3.27	3.74	3.64	4.01	4.15	3.58	3.59	3.07
	SD	2.34	1.14	1.71	.99	1.72	2.38	1.22	1.40	1.31
Time of Solution	M	73.71	100.50	107.38	76.75	99.21	141.31	144.85	161.04	153.07
	SD	20.67	33.59	29.31	21.42	28.02	87.25	37.91	51.35	106.59
Time of 1st Solution	M	56.83	57.15	52.76	41.89	51.75	73.88	70.08	77.10	74.37
	SD	40.11	28.07	21.94	15.66	24.65	40.03	25.48	48.47	32.72
Inter-test Time	M	13.38	15.95	21.05	13.57	13.33	15.45	19.41	26.16	23.34
	SD	8.39	6.05	7.28	4.36	5.03	8.89	10.19	11.55	11.79
Inter-Response Time	M	10.96	10.23	12.32	7.88	8.37	9.44	12.62	15.69	15.30
	SD	6.17	3.57	3.36	2.09	2.75	5.36	5.27	7.25	8.19
Number of Repeated Tests	M	.45	1.07	.80	.67	1.69	3.32	2.41	2.15	1.63
	SD	.59	1.01	.59	.56	1.70	2.17	2.58	2.50	1.31
Number of Repeated Solutions	M	.88	.48	.86	.27	.68	1.25	.72	.72	.27
	SD	2.22	.49	.71	.39	.80	1.39	.72	.68	.28

TABLE 4

Summary ANOVA of Mean Times-To-Solution During the Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	19217.02	2	9608.51	3.24*	.04
Functional Fidelity Level	58674.57	2	29337.29	9.88***	.0001
Physical x Functional	9907.46	4	2476.87	.83	.51
Error	240398.83	81	2967.89		

* p<.05
*** p<.001

TABLE 5

Summary ANOVA of Mean Times-to-First-Solution During the Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	1730.60	2	865.30	.83	.44
Functional Fidelity Level	6584.40	2	3292.20	3.15*	.05
Physical x Functional	4006.02	4	1001.50	.96	.43
Error	84624.27	81	1044.74		

* p<.05

TABLE 6

Summary ANOVA of Mean Inter-test Times During the Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	315.37	2	157.68	2.15	.12
Functional Fidelity Level	1236.93	2	618.46	8.44***	.001
Physical x Functional	246.82	4	61.70	.84	.50
Error	5935.89	81	73.28		

*** $p<.001$

TABLE 7

Summary ANOVA of Mean Inter-response Times During the Transfer Phase for the Various Simulator Training Groups

SOURCE	SS	df	MS	F	Prob.
Physical Fidelity Level	52.19	2	26.10	.94	.40
Functional Fidelity Level	537.98	2	268.99	9.68***	.0002
Physical x Functional	38.85	4	9.71	.35	.84
Error	2251.83	81	27.80		

*** p<.001

TABLE 8

Summary ANOVA of Mean Repeated Tests During the Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	8.35	2	4.17	1.55	.22
Functional Fidelity Level	29.48	2	14.74	5.48**	.01
Physical x Functional	32.32	4	8.08	3.00*	.02
Error	217.89	81	2.69		

* p<.05
*** p<.01

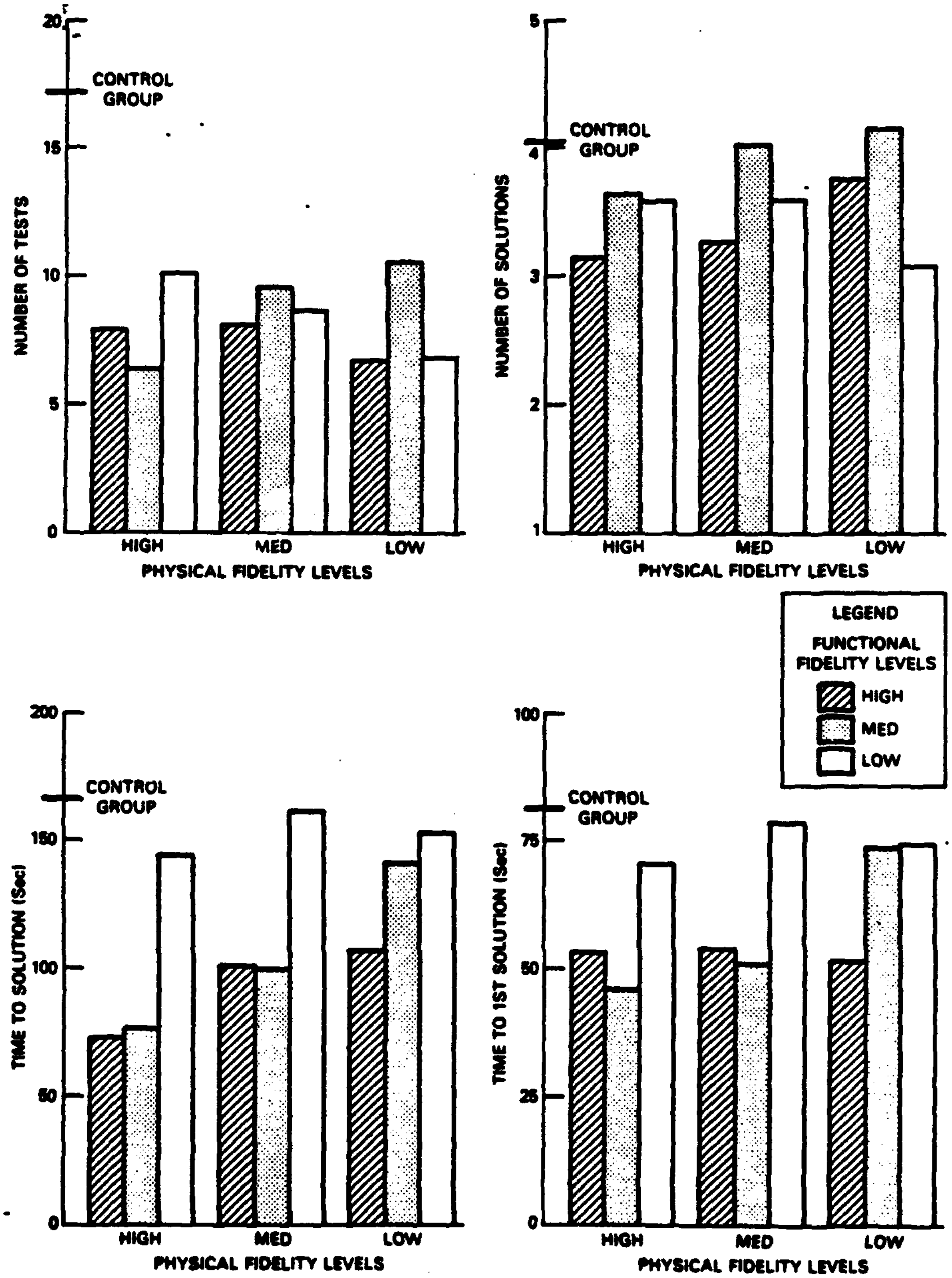

Figure 1. Mean performance of training groups as measured by number of tests, number of solutions, time to correct solution and time to 1st solution.

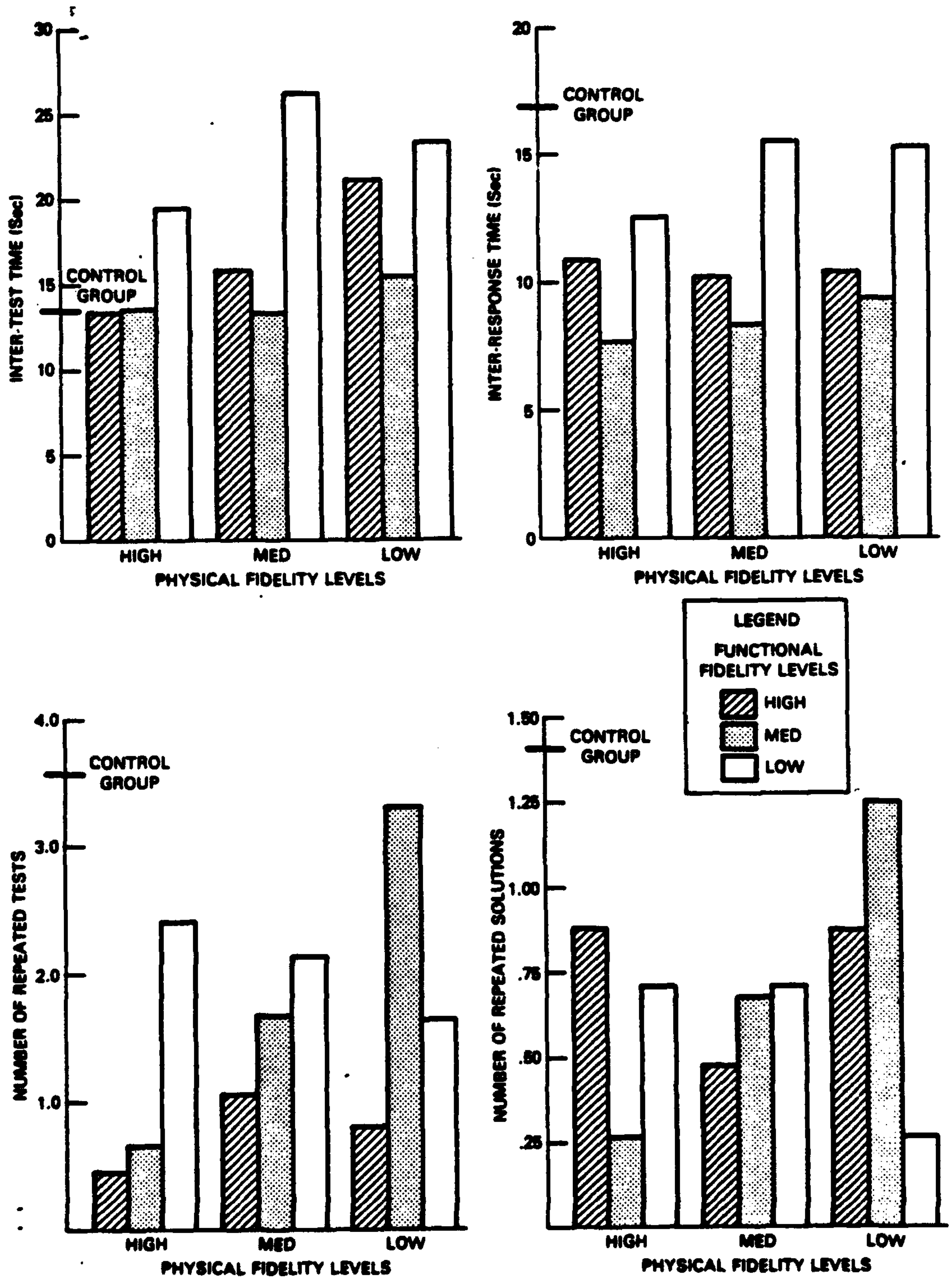

Figure 2. Mean performance of training groups as measured by inter-test time, inter-response time, number of repeated tests, and number of repeated solutions.

functional fidelity training produced superior performance.

As for interactions, only one between physical and functional fidelity on the number of repeated tests was noted $F(4,81) = 3.00$, $p < .02$. That is, fewer repeated tests were made by persons trained with high physical/high functional and medium physical/high functional simulators than for persons trained in other groups. When physical fidelity was degraded, groups trained with high functional fidelity performed especially well. The largest number of repeated tests were made by groups trained with medium functional and low physical fidelity. It would appear then, that training with relatively low levels of physical fidelity, and moderate functional fidelity may not provide enough information for subjects. In this case, subjects may not have realized that they were repeating tests because they did not understand the physical layout of the equipment in the transfer test.

As can be seen in Figures 1 and 2, the level of functional fidelity had a strong effect on performance, as measured by a number of dependent variables. In general, decreasing levels of functional fidelity were associated with longer solution times (Figure 1c) and longer intervals between actions (Figures 2a and 2b). This, of course, makes sense if one considers that with more information, problem solving may have become easier, and therefore, quicker.

The fact that lower levels of physical fidelity were also associated with longer solution times is less clear, however. One reason may be that subjects trained on the lower fidelity devices required more orientation time to the reference system

during transfer. This seems quite likely when one considers the visual disparity between the medium and low physical simulators and reference system.

The fact that temporal measures were especially sensitive to the fidelity manipulations was not surprising, and confirms similar findings in numerous field experiments involving training manipulations of various kinds (Bresnard and Briggs, 1956; Parker and Depaul, 1967; Rigney et al., 1978; Unger et al., 1984).

One of the more interesting and potentially important findings, however, was the significant interaction between fidelity levels with respect to the number of repeated tests. The interaction emphasizes the fact that physical and functional features in equipment are indeed related, and should not be dealt with in isolation. As for the interaction itself, Figure 2c reveals that the effect of decreasing physical similarity produced opposite effects in the medium and low functional groups. That is, for subjects trained in low functional conditions, there was a tendency to repeat fewer tests during transfer as physical similarity decreased. Subjects trained with medium functional fidelity, on the other hand, tended to repeat more and more tests with increasing dissimilarity between the training and transfer devices.

In summary, a two general conclusions seem warranted by the results of the ANOVA analyses. First, it is clear that temporal performance measures were most sensitive to fidelity manipulations since statistical significance was found on all four temporal measures. Especially sensitive was the total-time-to-solution variable which was found to be significantly

affected by variations in both functional and physical fidelity. Number of tests, number of solution attempts, and number of repeated solutions, on the other hand, seemed to operate quite independently of fidelity manipulations. Second, evidence was found to suggest that simulator fidelity should not be considered a single, uniform concept, but a multi-dimensional one consisting of at least a physical and functional component. Indeed, it was noted that each of these components often produced different effects depending on the dependent measure.

If reliable, this latter finding would appear to have far-reaching implications for those involved in training device design since it emphasizes the importance of carefully determining the critical aspects of tasks to be trained in order to ensure an appropriate mix of physical and functional features in the training device under consideration.

A.3.2 Post Hoc Comparisons of Control and Training Group Data

Since it was important to know whether performance on the criterion task was affected by the various training procedures, performance data from each of the nine simulator conditions was compared with that of the no-training control group. The results of these comparisons are shown in Table 9.

As can be seen in Figures 1 and 2, training group performances were generally superior to that of the no-training control group, although these differences did not always reach statistical significance as shown in Table 9. That is, subjects across training conditions generally attempted fewer solutions, solved problems quicker, and repeated fewer tests and solutions than their no-training counterparts. However, in view of the

TABLE 9

Results of Dunnett's Test Comparing the Performance of Training Groups with That of a No-Training Group on Eight Dependent Variables

			Physical Fidelity		
			High	Medium	Low
	High	Tests	-	-	p<.05
		Solutions	-	-	-
		Time-to-Solution	p<.01	-	-
		Time-To First-Solution	-	-	-
		Inter-Test Time	-	-	-
		Inter-Response Time	-	p<.05	-
		Test Repeats	p<.01	p<.01	p<.01
		Solution Repeats	-	-	-
Functional Fidelity	Medium	Tests	p<.01	-	-
		Solutions	-	-	-
		Time-To-Solution	p<.01	-	-
		Time-To-First Solution	p<.05	-	-
		Inter-Test Time	-	-	-
		Inter-Response Time	p<.01	p<.01	p<.05
		Test Repeats	p<.01	-	-
		Solution Repeats	-	-	-
	Low	Tests	-	-	p<.05
		Solutions	-	-	-
		Time-To-Solution	-	-	-
		Time-To-First Solution	-	-	-
		Inter-Test Time	-	p<.05	-
		Inter-Response Time	-	-	-
		Test Repeats	-	-	-
		Solution Repeats	-	-	-

fact that many of the performance differences were not statistically significant, these data suggest that the criterion task may not have been sufficiently difficult. Despite this, however, the effects of the training manipulations were powerful enough to significantly affect performance.

A.3.3 Individual Difference Correlates of Performance

In order to reduce the number of dependent variables for the individual difference analyses, the eight dependent variables were factor analyzed (principal factoring with iteration, Varimax rotation) with three factors emerging. The rotated factor matrix in Table 10 shows the emergence of three factors (time, tests and solutions). These three factors are parallel to those found by Henneman and Rouse (1984) of time, inefficiency, and errors.

Pearson correlations. Pearson correlations were computed between each individual difference measure and each of the three performance factors. Significant correlations appear in Table 11. Positive correlations were noted between time and GRE-analytic, VPI-intellectual and VPI-masculinity scores. In other words, persons with high intellectual or masculine interests, or high analytic ability, generally took longer to solve the criterion problems. With respect to the test factor, persons high on either mechanical or analytic ability, or with low enterprising interests, tended to make fewer tests. Field independent persons with high analytic abilities made fewer incorrect solutions. While this pattern of results (negative correlations with tests and solution factors; positive correlation with time) appears paradoxical, an explanation may be found in subjects'

TABLE 10

Varimax-Rotated Factor Matrix

	Factor 1 "Time"	Factor 2 "Tests"	Factor 3 "Solutions"
TTIME	.740 *	.509	.046
TFS	.674 *	.281	-.402
ITT	.831 *	-.312	-.015
IRT	.941 *	-.103	-.006
NTEST	-.226	.854 *	.269
TREP	-.003	.818*	.131
NSOL	-.026	.116	.918*
SOLREP	.068	.275	.789*

* loadings on each factor

TABLE 11

Correlations Between Individual Difference Measures and the Three Performance Factors

	Performance Factors		
	Time	Tests	Attempted Solutions
GEFT			-.24*
BMC		-.30**	
GRE-A	.23*	-.22*	-.24*
LOC			
VPI-Realistic			
VPI-Intellectual	.23*		
VPI-Social			
VPI-Conventional			
VPI-Enterprising		.26*	
VPI-Masculinity	.19*		
Sex			
Age			

*p<.05

**p<.01

troubleshooting strategies. That is, analytic individuals may have taken longer to reach solutions since they may have approached problems more thoughtfully than non-analytic subjects. Indeed, the latter subjects may have relied mainly on multiple testing rather than an understanding of the system.

<u>Interaction of individual differences and training conditions</u>. To determine the predictive strength of physical and functional fidelity, the various individual difference measures, and the individual difference x fidelity interactions, 11 multiple regressions were computed for each of the three performance factors. Due to sample size limitations, it was not feasible to include the large number of individual difference measures and their interactions with fidelity in a single equation. In accordance with standard procedure for moderated regression equations (Ghiselli, Campbell, & Zedeck, 1981), physical fidelity level, functional fidelity level, and a given individual difference measure were forced into the equation prior to their respective interactions. This procedure allows the testing for significant interactive effects over and above the contributions made by the main effects of individual differences and fidelity. The results of these multiple regressions appear in Table 12. It is clear that the time factor was the most sensitive to fidelity manipulations. The subject's functional and physical fidelity training condition predicted a significant amount of variance in each regression for this factor. This is consistent with the results of the ANOVA analyses reported above.

No individual difference variables significantly predicted the time factor, and only two interactions involving an individual

TABLE 12

Predictors of Performance on Three Varimax-Rotated Factors: Time, Tests and Solutions.

	Time Factor	Test Factor	Solution Factor
GEFT	Func**, Phys*		GEFT*
BMC	Func**, Phys*	BMC*	
GRE	Func**, Phys*		GRE*, GRExFxP*
LOC	Func**, Phys*		
VPI-R	Func**, Phys*		
VPI-I	Func**, Phys* VPI-IxF*		
VPI-S	Func**, Phys*	VPI-SxF*	VPI-SxP*
VPI-C	Func**, Phys*		
VPI-M	Func**, Phys*		VPI-MxP*
AGE	Func**, Phys* AgexFxP*		
GENDER	Func**, Phys*		GenderxF*

* $p<.05$
** $p<.01$

difference measure were significant. The VPI intellectual scale interacted with functional fidelity such that individuals high in intellectual interests took more time under the low functional fidelity condition than did their low intellectual interest counterparts. A three-way interaction between age, functional and physical fidelity indicated that older individuals took significently longer times to solution when trained under low functional, low physical conditions.

As for the test factor, only two significant predictors were noted. Individuals with high scores on mechanical comprehension attempted fewer tests. A significant interaction between VPI social and functional fidelity was also found. That is, individuals high in social interests attempted more tests under low and medium functional fidelity conditions but not in high functional fidelity conditions.

Two main effects and four interactions were found on the attempted solutions factor. Individuals high on field independence or analytic ability attempted fewer solutions. A significant three-way interaction was found between GRE-A, physical, and functional fidelity. In this case, individuals high in analytic ability attempted fewer solutions after training in high physical, high functional fidelity conditions than their low analytic counterparts.

Gender and functional fidelity also interacted significantly indicating that males attempted fewer solutions after training in low functional conditions whereas females attempted fewer solutions after training in high functional conditions.

The two-way interactions of VPI-Social x physical fidelity

and VPI-Masculinity x physical fidelity showed similar patterns. Individuals with low social or high masculine interests tried fewer solutions under low physical training conditions, whereas high social or low masculine individuals tried fewer solutions under medium and high physical training conditions. These parallel findings are due to the negative correlation found in this study between social and masculine interest scores on the VPI, $r=-.57$. An explanation may be that those with masculine interests gain considerable knowledge of the system under the low physical condition which essentially provides them with a schematic diagram. However, those who are low on masculine interests may fail to make the connection between the schematic and the reference system and require high physical fidelity for optimum performance.

The number of significant interactions do indicate a pattern of results supportive of an adaptive training model, i.e., optimum training may require different modes of training for individuals with different abilities and aptitudes (Rouse, 1984). Cross-validations are required for more definitive conclusions.

A.4 Summary and Conclusions

The results of this study lead to the following conclusions:

1. In the present task, the physical and functional aspects of fidelity were separately manipulated and yielded differential effects. This is important since previous research has not systematically investigated these two aspects of fidelity.
2. The effect of functional fidelity was a very potent determinant of performance. Decreasing levels of

functional fidelity were generally associated with longer solution times and longer inter-response times.

3. The manipulation of the physical aspect of fidelity also produced a significant performance effect, as measured by time to solution. Persons trained on simulators with lower physical fidelity took longer to reach the correct solution.

4. An interaction of physical and functional fidelity was only found on the number of repeated tests. The effect of decreasing the amount of physical fidelity during training produced different performance levels depending on the level of functional fidelity. Subjects repeated more tests in the medium functional conditions as physical fidelity decreased while subjects in the low functional conditions repeated fewer tests as physical fidelity decreased.

5. Although subjects trained on simulators performed better than subjects in the no-training control group, the data suggest that criterion problems should be of higher difficulty.

6. Factor analysis of the eight dependent variables yielded three performance factors: time, tests, and solutions. These factors are similar to those found by Henneman and Rouse (1984) in their studies of human fault detection.

7. Persons with high analytic abilities (GRE-A) took longer to solve criterion problems but required fewer tests and attempted fewer incorrect solutions. This may reflect more thoughtful approach rather than relying on simple multiple testing.

8. The time factor was the most sensitive to fidelity manipulations. Subjects' training condition was the most potent predictor of performance on this factor.

9. Mechanical comprehension level was the strongest predictor of performance on the tests factor. Those subjects with high mechanical comprehension required fewer tests to reach a solution. This probably indicates that these subjects had a more complete understanding of the system than other trainees.

10. Several interesting interactions predicted performance on the attempted solutions factor. Both the subject's masculine and social interests were found to interact with physical fidelity. Analytical ability interacted with both aspects of fidelity while gender only interacted with the functional component.

B. Transfer Study II

B.1 Purpose of Experiment II

Since evidence was found in Experiment I to sugggest that the criterion problems used during the transfer phase may have been too easy, a second experiment, identical to the first but with more difficult problems during transfer, was run. A general expectation was that greater differences between the performances of the various simulator training groups and a no-training control group would result. It was also felt that a replication would allow us to check the reliability of some of the results noted in Experiment I and, at the same time, take a preliminary look at the effects of problem difficulty on transfer performance.

B.2 Method

B.2.1 Subjects

One hundred college undergraduates (50 males and 50 females) ranging in age from 16 to 40 served as paid subjects ($5.00 per hour). They were drawn primarily from introductory psychology courses at George Mason University.

B.2.2 Materials

Apparatus and individual difference measures were identical to those used in Experiment I.

B.2.3 Design and Procedure

Except for the use of a more difficult set of criterion problems during the transfer phase, the experimental design and procedure were identical to those used in Experiment I. Selection of problems for the present experiment was based on previous baseline data with the reference system and the choice of problems which were more difficult than those in Experiment I in terms of number of tests, solution attempts and/or times-to-solution.

B.3 Results and Discussion

B.3.1 ANOVAS on Dependent Variables

Three x three factorial analyses of variance comparing the transfer performance of the groups training in the various fidelity conditions were performed on each of the eight dependent variables. Means and standard deviations for the various groups are shown in Table 13. Summary analysis of variance tables containing significant factors are presented in Tables 14 to 17.

As was true in Experiment I, main effects of functional fidelity manipulations were found to be significant for a number

TABLE 13

Means and Standard Deviations for Dependent Variables on Transfer Test (Experiment II)

Variables		Training Conditions (physical fidelity/functional fidelity)								
		HH	MH	LH	HM	MM	LM	HL	ML	LL
Number of Tests	M	12.68	15.39	16.68	14.48	17.00	15.27	17.59	17.24	13.92
	SD	6.16	6.12	7.81	4.54	7.29	7.78	6.84	12.90	6.94
Number of Solutions	M	6.87	6.14	6.99	8.14	8.52	7.55	7.30	10.50	5.72
	SD	3.50	3.52	5.13	1.99	2.25	3.27	3.53	5.49	2.47
Time of Solutions	M	164.50	203.03	237.85	171.77	180.50	220.61	320.34	298.63	244.56
	SD	76.47	73.19	92.67	65.08	105.64	125.49	78.85	214.49	81.35
Time of 1st Solution	M	48.92	78.95	100.55	57.43	65.30	85.30	122.10	129.35	99.27
	SD	37.17	50.53	60.14	46.74	56.22	64.01	47.74	157.20	50.48
Inter-test Time	M	18.14	15.61	20.99	14.22	10.85	14.97	22.08	22.54	22.73
	SD	11.07	7.64	12.74	8.27	5.25	5.91	7.27	19.61	11.52
Inter-Response Time	M	8.94	16.40	10.70	8.20	6.93	8.75	13.76	11.05	19.63
	SD	5.33	14.09	3.66	4.98	3.17	3.97	5.04	6.46	25.46
Number of Repeated Tests	M	1.67	2.41	3.55	1.37	2.30	2.64	4.15	3.12	2.86
	SD	1.79	2.18	3.87	1.09	2.54	1.82	2.28	3.16	1.92
Number of Repeated Solutions	M	1.45	1.30	1.63	1.33	1.37	2.22	1.88	2.52	1.69
	SD	2.11	1.21	1.96	.86	.94	1.44	1.63	1.89	1.59

TABLE 14

Summary ANOVA of Mean Times-to-Solution During the Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	3602.37	2	1801.19	.15	.86
Functional Fidelity Level	169084.94	2	84542.47	6.93**	.002
Physical x Functional	67343.61	4	16835.90	1.38	.25
Error	987970.81	81	12197.17		

** $p<.01$

TABLE 15

Summary ANOVA of Mean Times-to-First-Solution During the Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	5980.5	2	2990.25	.58	.56
Functional Fidelity Level	39707.99	2	19853.99	3.84*	.03
Physical x Functional	16526.39	4	4131.60	.80	.53
Error	418813.48	81	5170.54		

* $p<.05$.

TABLE 16

Summary ANOVA of Mean Inter-test Times During Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	157.04	2	78.52	.68	.511
Functional Fidelity Level	12454.54	2	622.77	5.37**	.01
Physical x Functional	86.21	4	21.55	.19	.95
Error	9400.48	81	116.06		

** $p<.01$

TABLE 17

Summary ANOVA of Mean Inter-response Times During Transfer Phase for the Various Simulator Training Groups

Source	SS	df	MS	F	Prob.
Physical Fidelity Level	112.02	2	56.01	.50	.61
Functional Fidelity Level	712.44	2	356.22	3.19*	.05
Physical x Functional	594.07	4	148.52	1.33	.27
Error	9055.90	81	111.80		

* $p<.05$

of variables, viz., total-time-to-solution $F(2,81)=6.93$, $p<.01$; time-to-first-solution $F(2,81)=3.84$, $p<.03$; inter-test time $F(2,81)=5.37$, $p<.01$; and inter-response time $F(2,81)=3.19$, $p<.05$. It should be noted that apart from repeated tests, the same dependent variables which were found to be sensitive to functional fidelity manipulations in Experiment I were sensitive in the present data. Thus, once again, more feedback during training was generally associated with better (at least quicker!) performance in terms of the temporal dependent measures. Subjects experiencing lower levels of feedback during training generally took longer to reach correct solutions and more time between actions.

Unlike Experiment I, however, physical fidelity manipulations did not significantly affect transfer performance. It will be recalled that in Experiment I, both a significant main effect for physical fidelity on the total-time-to-solution variable as well as a significant interaction between physical and functional fidelity on repeated tests were found. No such effects or interactions were revealed in the present data.

Thus, as was the case in Experiment I, temporal dependent variables such as time-to-solution, time-to-first-solution, inter-test time, etc. were found to be especially sensitive to functional fidelity manipulations. Number of tests, number of solution attempts, number of repeated tests and number of repeated solutions, on the other hand, were shown to be relatively insensitive. The most surprising difference between the results of the present experiment and Experiment I was the failure to find any evidence that physical fidelity had any significant affect on transfer performance.

B.3.2 Post Hoc Comparisions of Control and Training Group Data

Since a major aim of the present experiment was to see whether more difficult criterion test problems would increase the differences between the performances of simulator-trained and no-training control subjects, a number of post hoc comparisons between control and training group data were performed. The results of these tests are given in Table 18.

As can be seen, subjects receiving simulator training performed significantly better than control subjects on a number of dependent variables. That is, they generally made fewer tests, solved problems quicker, repeated fewer tests and solutions and, in some cases, were more apt to offer problem solutions earlier than their no-training counterparts. A comparison of Tables 3 and 13 also makes it clear that those receiving some form of training benefited most when criterion problems were made more difficult. That is, whereas only 15 significant differences between training and no-training group data were noted in Experiment I, 26 emerged in Experiment II.

B.3.3. Individual Difference Correlates of Performance

As was the case in Experiment I, the eight dependent variables were factor analyzed (principal factoring with iteration, varimax rotation) in order to reduce their number for individual difference analyses. The rotated factor matrix in Table 19 shows the emergence of three factors: total time and tests, inter-test time, and solution attempts. It is interesting to note that the first two factors differ from those in Experiment I in that number of tests attempted and tests repeated loaded on the same factor as total time, while inter-test and inter-response

TABLE 18

Results of Dunnett's Test Comparing the Performance of Training-Groups with That of a No-Training Group on Eight Dependent Variables (Experiment II)

			Physical Fidelity		
			High	Medium	Low
	High	Tests	p<.005	p<.025	p<.05
		Solutions	-	-	-
		Time-to-Solution	p<.005	p<.005	p<.005
		Time-to-First-Solution	p<.025	-	-
		Inter-Test Time	-	-	-
		Inter-Response Time	-	-	-
		Test Repeats	p<.005	p<.005	p<.05
		Solution Repeats	-	-	-
Functional Fidelity	Medium	Tests	p<.01	p<.05	p<.01
		Solutions	-	-	-
		Time-to-Solution	p<.005	p<.005	p<.005
		Time-to-First-Solution	p<.05	-	-
		Inter-Test Time	-	-	-
		Inter-Response Time	-	-	-
		Test Repeats	p<.005	p<.005	p<.005
		Solution Repeats	-	-	-
	Low	Tests	p<.05	p<.05	p<.005
		Solutions	-	-	-
		Time-to-Solution	-	-	p<.01
		Time-to-First-Solution	-	-	-
		Inter-Test Time	-	-	-
		Inter-Response Time	-	-	-
		Test Repeats	-	p<.025	p<.01
		Solution Repeats	-	-	-

TABLE 19

Varimax-Rotated Factor Matrix (Experiment II)

	Factor 1 "Test + Total Time"	Factor 2 "Inter-test Time"	Factor 3 "Attempted Solutions"
Number of Tests	.847	-.391	.158
Repeated Tests	.703	-.116	.011
Total Time	.878	.427	.159
Time-to-First-Solution	.683	.456	-.190
Inter-test Time	-.123	.927	.006
Inter-response Time	.073	.561	-.168
Number of Solutions	-.074	-.226	.974
Repeated Solutions	.128	-.007	.687

time emerged as a separate factor. The different factor structure was no doubt due to the increased problem difficulty in Experiment II, as demonstrated by the analyses of variance results. Clearly more difficult problems require more tests which, as inter-test time remains the same, results in longer time taken to solve the problem. Hence the correlation between tests and total time is much stronger in Experiment II ($r=.59$, $p<.001$) than in Experiment I ($r=.27$, $p<.005$). Unfortunately, the fact that the factor structure was different prevented combining data across the two experiments so that individual difference effects might be investigated.

Pearson correlations. Pearson correlations were computed between each individual difference measure and each of the three performance factors found in Experiment II. These correlations appear in Table 20. With respect to the tests and total time factor, persons with either high masculine or low social interests tended to take longer and/or make more tests. Also, the positive correlation between GRE-analytic and inter-test time indicated that persons with strong analytic abilities deliberated longer between tests. Moreover, persons with masculine interests attempted fewer incorrect solutions.

Separate correlation matrices for both Experiment I and II were also computed between the individual difference measures and the eight dependent variables. Only 3 significant relationships found in Experiment I were replicated in Experiment II:

1. Persons with strong analytic (GRE-A) ability tended to make fewer tests in each experiment ($r=-.19$, $r=-.18$ respectively).

TABLE 20

Significant Predictions in Moderated Regressions for each of the Performance Factors (Experiment II)

Individual Difference Measures	Performance Factors: Test + Total Time	Inter-test Time	Solutions Attempted
GEFT	GEFT x Func*	Func*	
BMC	BMC x Phys*, BMC x F x P*	Func*	
GRE-A	GRE x Phy*, Phy x Fun*	Func*	
LOC	LOC x F x P*	Func*	
VPI-Realistic	VPI-R x F x P*	Func*, VPI-R x F	
VPI-Intellectual	VPI-I x F x P*	Func*, VPI-I x F	
VPI-Social	VPI-S*,	Func*	
VPI-Conventional	Phy x Func*	Func*	
VPI-Masculinity	VPI-M***, Func*	Func*	

* p<.05
** p<.01
*** p<.001

2. Persons with high intellectual interests (VPI-I) had longer inter-response times (r=.25, r=.27).

3. Persons with high social interests (VPI-S) tended to repeat fewer tests (r=-.20, r=-.23).

The first of these findings is not surprising; however the other two require some explanation. Persons with intellectual interests are likely to be more reflective and thoughtful, thus deliberating longer prior to making their next response. Those with social interest superficially would appear to be at a disadvantage on a mechanical task that requires no social contact. However, it is possible that their distaste for the activity may have led them to adopt a sequential strategy (testing each relay in numerical order). Such a strategy may have made it easy to remember which relays had been tested, thus lowering the probability of repeating a test.

<u>Interaction of individual differences and training conditions</u>

In order to determine the relative strength of the fidelity manipulations and the various individual difference measures, nine multiple regressions were computed for each of the three performance factors. For each of these hierarchical stepwise multiple regressions, physical fidelity level, functional fidelity level, and a given individual difference measure were forced into the equation prior to interactions. The results of these multiple regressions appear in Table 21. The significant effect for functional fidelity in predicting the inter-test time factor is consistent with the previous Experiment II ANOVA analyses as is the nonsignificant relationship between functional fidelity and the attempted solution factor. The failure of functional

TABLE 21

Correlations Between Individual Difference Measures and the Three Performance Factors (Experiment II)

	Performance Factors		
	Test + Time	Inter-Test	Attempted Solutions
GEFT			
BMC			
GRE-A		.19*	
LOC			
VPI-Realistic			
VPI-Intellectual			
VPI-Social	-.26**		
VPI-Conventional			
VPI-Enterprising			
VPI-A			
VPI-SC			
VPI-Masculinity	.29**		-.20*
VPI-ST			
VPI-F			
VPI-AQ			

* p<.05
** p<.01

fidelity to predict the test/total time factor also parallels the ANOVA results which found no significant effects for two (tests, repeated tests) of the four dependent variables that load highly on the factor.

Although there were significant effects of masculine and social interests for the tests/total time factor, these have been discussed previously in the correlational analyses. However, two of the significant two-way interactions that involved individual differences are particularly striking. Both mechanical and analytic ability interacted with physical fidelity such that individuals strong in those abilities did worst (longer total time/more tests) under high physical fidelity conditions whereas their low ability counterparts did best under the same conditions. Moreover, while there were minimal performance differences between high and low analytic individuals under high physical fidelity conditions, high analytic subjects considerably outperformed low analytic individuals under low physical fidelity conditions. Thus high physical fidelity appeared to aid those equipped with lesser abilities for the task whereas for those with high abilities a more cognitive schema (low physical fidelity) facilitates. These interactions may well explain the failure to find significant main effects of physical fidelity. In addition, they have substantial applied implications for training on troubleshooting tasks.

Both realistic and intellectual interests showed the same pattern in their interactions with functional fidelity and in predicting inter-test time. Individuals with strong realistic or intellectual interests in the low functional condition were

particularly inclined toward lengthy inter-test times. The lack of feedback apparently forced them to be particularly deliberative in moving to the next test. The significant interaction between field independence and functional fidelity is more difficult to explain. In general, field independent people took less time and made fewer tests. However, under medium functional fidelity conditions field dependent individuals performed somewhat better.

Regretably, none of the three-way interactions appeared interpretable. However, the general pattern of interactions between individuals differences and training conditions would seem to support an adaptive training model, i.e., different modes of training may be suited for different individuals. This is particularly true of the interactions involving physical fidelity. Clearly these data point to the use of low physical fidelity (at presumably a lower cost) simulators with a high-ability trainee group, with high physical fidelity simulators limited to low-ability trainees.

B.3.4 Analysis of Combined Data from Experiments I and II

Although studies I and II were separate experiments, the fact that the only major difference between them was the use of a more difficult problem set in Experiment II provided an opportunity to examine, in a preliminary way, the effects of problem difficulty and its interactions with physical and functional fidelity on transfer performance. To this end, the data from the two experiments were combined and analyzed by means of 2 x 3 x 3 ANOVAs (problem difficulty x physical fidelity level x functional fidelity level) on each of the eight dependent variables. The results of these analyses are summarized in

Tables 22 to 29.

As can be seen, problem difficulty had a very strong and significant effect on transfer performance as measured by six of the eight dependent variables. Indeed, only the inter-test and inter-response time variables failed to be affected. Thus, regardless of training group, subjects in Experiment II generally made more tests, offered more incorrect solutions, took longer to solve problems, waited longer to offer their first solution, and repeated more tests and solutions than subjects in Experiment I.

As for physical and functional fidelity, not surprisingly, significant main effects for functional fidelity on the time-to-solution, time-to-first-solution, inter-test time and inter-response time variables were found once again. In addition, as was the case in Experiment I, a significant main effect for physical fidelity as well as an interaction between physical fidelity and functional fidelity on the repeated tests variable was noted.

The most interesting discovery, however, was the failure to find any significant interactions between problem difficulty and either physical or functional fidelity. It would appear then, that despite dramatic effects on performance, problem difficulty operated rather independently of physical and functional manipulations during training. That is, subjects in the various training groups responded rather similarly to more difficult problems during the transfer phase and simply made more tests and solution attempts, took longer to solve problems, etc. If reliable, this result would seem to have important implications for training and training-device design, since it suggests that

TABLE 22

Summary ANOVA of Mean Number of Tests During the Transfer Phase For the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	2364.15	1	2364.15	60.52***	.001
Physical Fidelity Level	46.81	2	23.41	.60	.55
Functional Fidelity Level	46.79	2	23.40	.60	.551
Difficulty x Physical	6.61	2	3.31	.09	.92
Difficulty x Functional	7.65	2	3.8	.10	.91
Physical x Functional	205.99	4	51.50	1.32	.27
Difficulty x Physical x Functional	99.72	4	24.93	.64	.64
Error	6328.40	162	39.06		

*** $p<.001$

TABLE 23

Summary ANOVA of Mean Number of Solutions During the Transfer Phase for the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	700.77	1	700.77	87.70***	.001
Physical Fidelity Level	20.26	2	10.13	1.27	.28
Functional Fidelity Level	28.99	2	14.49	1.81	.17
Difficulty x Physical	20.64	2	10.32	1.29	.28
Difficulty x Functional	10.79	2	5.40	.68	.51
Physical x Functional	57.55	4	14.39	1.80	.13
Difficulty x Physical x Functional	34.17	4	8.54	1.07	.37
Error	1294.47	162	7.99		

*** $p<.001$

TABLE 24

Summary ANOVA of Times-to-Solution During the Transfer Phase for the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	537881.73	1	537881.73	70.94***	.001
Physical Fidelity Level	19709.58	2	9854.79	1.3	.28
Functional Fidelity Level	209225.84	2	104612.92	13.80***	.001
Difficulty x Physical	3112.56	2	1556.28	.21	.82
Difficulty x Functional	18534.84	2	9267.42	1.22	.30
Physical x Functional	57774.95	4	14443.74	1.91	.11
Difficulty x Physical x Functional	19472.62	4	4868.16	.64	.63
Error	1228336.36	162	7582.32		

*** $p<.001$

TABLE 25

Summary ANOVA of Mean Times-to-First Solution During the Transfer Phase for the Various Transfer Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	29734.64	1	29734.64	9.57**	.002
Physical Fidelity Level	6937.39	2	3468.70	1.12	.33
Functional Fidelity Level	39145.98	2	19572.99	6.30**	.002
Difficulty x Physical	773.71	2	386.86	.124	.88
Difficulty x Functional	7146.42	2	3573.21	1.15	.32
Physical x Functional	11020.80	4	2755.20	.89	.47
Difficulty x Physical x Functional	9511.6	4	2377.9	.77	.55
Error	503437.75	162	3107.64		

** p<.01

TABLE 26

Summary ANOVA of Mean Inter-test Times During the Transfer Phase for the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	.13	1	.13	.001	.97
Physical Fidelity Level	292.17	2	146.08	1.54	.22
Functional Fidelity Level	2437.84	2	1218.92	12.88**	.001
Difficulty x Physical	180.25	2	90.12	.95	.39
Difficulty x Functional	44.63	2	22.31	.24	.79
Physical x Functional	306.59	4	76.65	.81	.52
Difficulty x Physical x Functional	26.43	4	6.61	.07	.99
Error	15336.37	162	94.67		

*** p<.001

TABLE 27

Summary ANOVA of Mean Inter-response Times During the Transfer Phase for the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	1.34	1	1.34	.02	.89
Physical Fidelity Level	158.25	2	79.13	1.13	.32
Functional Fidelity Level	1234.47	2	617.23	8.84***	.001
Difficulty x Physical	5.96	2	2.98	.04	.96
Difficulty x Functional	15.96	2	7.98	.11	.89
Physical x Functional	211.01	4	52.75	.76	.56
Difficulty x Physical x Functional	421.92	4	105.48	1.51	.20
Error	11307.73	162	69.80		

*** $p<.001$

TABLE 28

Summary ANOVA of Mean Number Repeated Tests During the Transfer Phase for the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	54.09	1	54.09	12.65***	.001
Physical Fidelity Level	13.83	2	6.92	1.62	.20
Functional Fidelity Level	35.41	2	17.70	4.14*	.02
Difficulty x Physical	.46	2	.23	.05	.95
Difficulty x Functional	19.24	2	9.62	2.25	.11
Physical x Functional	47.90	4	11.98	2.80*	.03
Difficulty x Physical x Functional	14.20	4	3.55	.83	.51
Error	692.52	162	4.28		

*   p<.05
*** p<.001

TABLE 29

Summary ANOVA of Mean Number of Solutions Repeated During the Transfer Phase for the Various Simulator Training Groups (Combined Data: Experiments I and II)

Source	SS	df	MF	F	Prob.
Problem Difficulty Level	47.86	1	47.86	27.19***	.001
Physical Fidelity Level	1.40	2	.80	.43	.64
Functional Fidelity Level	1.17	2	.58	.33	.72
Difficulty x Physical	.24	2	.12	.07	.93
Difficulty x Functional	4.46	2	2.23	1.27	.28
Physical x Functional	13.44	4	3.36	1.91	.11
Difficulty x Physical x Functional	1.36	4	.34	.19	.94
Error	285.18	162	1.76		

*** $p<.001$

designers may be able to manipulate fidelity in training devices without undue concern about the difficulty of subsequent "real world" problems. To be sure, however, such a conclusion it, at best, a provisional one, and must await additional supporting research.

B.4 Summary and Conclusions

The results of Experiment II may be summarized as follows:

1. Although functional fidelity manipulations during training produced a number of significant effects during transfer, unlike Experiment I, the physical fidelity of training simulators did not have any appreciable effect on subjects' performance. Also, no interactions between physical and functional fidelity were noted.

2. As expected, subjects given some form of simulator training performed significantly better than subjects receiving no training at all. Clearly, practice on problems during training helped simulator groups solve the more difficult problems in Experiment II more efficiently and effectively.

3. As was the case in Experiment I, temporal variables such as total-time-to-solution, time-to-first-solution, inter-test time and inter-response time were especially responsive to fidelity manipulations during training. This was particularly true for functional fidelity.

4. The more difficult criterion problems in Experiment II had a similar effect across simulator groups. That is, regardless of training condition, subjects took more time to solve problems, offered first solutions later, made more tests and solutions and repeated more tests and solutions than their

Experiment I counterparts. No particular training group seemed to have enjoyed a significant advantage.

5. Factor analysis of the eight dependent variables yielded a somewhat different set of performance factors than those found in Experiment I, viz. total time and tests, inter-test time and solution attempts. This different factor structure was apparently due to the more difficult criterion problems used in Experiment I.

6. Correlations between individual difference measures and dependent variables revealed that three significant relationships in Experiment I were replicated in Experiment II. That is, once again, persons with high analytic ability (GRE-A) tended to make relatively few tests; persons with strong intellectual interests (VPI-I) demonstrated relatively long inter-response times; and finally, persons with strong social interests (VPI-S) tended to repeat relatively few tests.

B.5 General Conclusions: Experiments I and II

Considered together, a number of broad conclusions may be drawn from the results of Experiments I and II.

1. Evidence was found to suggest that simulator fidelity should not be considered a single, uniform concept, but a multi-dimensional one consisting of at least a physical and functional component. Indeed, it was noted that each of these components sometimes produced different effects depending on the dependent measure used during transfer. Such a finding has far-reaching implications for training device design since it emphasizes the importance of carefully determining the critical aspects of tasks to-be-trained so that an appropriate mix of physical and functional

features may be incorporated in the training device under consideration.

2. With "cognitive" tasks like the present one, the use of training simulators differing along functional and, to a lesser extent, physical fidelity dimensions can lead to significant performance differences during subsequent transfer.

3. The fact that differential fidelity effect sometimes occurred depending on the dependent variable selected, suggests that the choice of criterion performance measures may be especially important in simulator design decisions.

4. Temporal variables like time-to-solution, time between tests, time between solution attempts, etc. seem to be unusually sensitive to fidelity manipulations. Number of times a subject repeats a particular test during problem solving may also be affected.

5. No evidence was found to suggest that there was any advantage to "actual equipment" training in the present experiments. If there was any benefit, it was a slight and nonsignificant one, and reflected a tendency for high physical/high functional subjects to solve problems a bit quicker and without repeating so many tests.

6. Evidence was found to suggest that the difficulty level of criterion problems during transfer affects performance in a rather general way, and quite independently of simulator fidelity levels experienced during training. Thus, at least with the difficulty levels examined in the present experiments, no particular training condition seemed to fare better than any other.

7. Certain individual difference variables seem to correlate reliably with particular performance characteristics such that the following generalizations are possible:

a. Persons with relatively high analytic abilities (GRE-A) make fewer tests than persons with lower abilities.

b. Subjects with strong intellectual (VPI-I) interests take more time between tests than less intellectually-inclined individuals.

c. Persons with strong social interests and preferences (VPI-S) tend to repeat fewer tests during problem solving than less socially-oriented individuals.

8. The fact that a number of significant interactions were noted between certain individual differences variables and training conditions suggests that the concept of adaptive training may be an especially useful one in training situations like the present one. Indeed, the data indicate that persons with rather low mechanical and analytic abilities may respond best to training conditions in which high physical fidelity is present, whereas individuals with high abilities may thrive under low physical fidelity conditions. A number of generalizations would appear to follow:

a. Certain abilities may predispose individuals to adopt particular problem-solving strategies. For example, high "analytic" persons may attack problems in a deliberative, logical way in an attempt to understand the system.

b. Certain training conditions facilitate particular problem-solving strategies. For example, low physical fidelity conditions may help by providing a "cognitive map" of the system which may facilitate the adoption of a logical strategy for aiding system understanding.

c. When both a subject predisposition and training condition are congruent with respect to adoption of a particular strategy, performance will be enhanced. However, when incogruent (e.g., low analytic individuals trained on low physical fidelity simulators), performance may be hindered.

VIII. Future Research and Recommendations

A. Training Systems Research Issues and Simulator Fidelity

In terms of future research directions, we believe that two general areas deserve further attention, viz., (1) additional research aimed at further specifying minimal fidelity levels in not only tasks and situations like those addressed in the present work, but in other task domains and with other task types as well, and (2) research designed to examine and account for the effects of fidelity-interactive variables on trainee learning and transfer performance.

With regard to the first area, we are inclined to agree with Baum, et al. (1982) that while future work should continue in the equipment maintenance and repair arena, task domains such as equipment operation and command and control (C^2) should also be addressed. Within each of the areas, procedural, perceptual-motor

and cognitive tasks also need to be examined, and at both the individual and team training level. A useful general research framework blending these three dimensions has been neatly laid out by Baum, et al. (1982).

However, as desirable as it might be to know which general levels of fidelity are sufficient for training certain skills and knowledges, such specifications will always be of limited value if they do not take into account the effects of such factors as individual differences, user-acceptance issues, task difficulty, stage of training, and a host of other "fidelity-interactive" variables. Unfortunately there is little in the way of guidance, apart from scattered generalizations, available to assist the training systems developer make proper and cogent decisions. As Hays and Singer (1983) have pointed out, "Systematic research is needed to determine how the relative values of each aspect of fidelity interact with other training systems variables to produce various degrees of transfer of training. (p. 16)" In this regard, we believe that research should be performed which centers not only on the general issue of device fidelity and its relationship to training effectiveness and transfer, but on other important training systems issues and variables as well. In terms of the 4-step model of the Instructional Systems Development Process (ISD) envisioned by Hays and Singer (1983), a number of these critical research issues can be identified. They include:

1. Training Needs Assessment Issues

   a. Trainee Characteristics
      (1) Enabling Skills
      (2) Trainee Strategies
      (3) Aptitudes
      (4) Abilities

   b. Training Strategies
      (1) Generic-versus specific-device training

2. Training Device Design Issues

   a. Simulator Fidelity
      (1) Physical/functional fidelity

   b. Instructional Features
      (1) Amount of training
      (2) Cost of actions/decisions
      (3) Instructional set/pretraining
      (4) Feedback

3. Training Device Implementation Issues

   a. User Acceptance
      (1) Trainer
      (2) Trainee

4. Training Effectiveness Assessment Issues

   a. Retention
      (1) Delays between simulator and AET training
      (2) Long-term Retention

B. <u>Training Needs Assessment Issues</u>

B.1 <u>Trainee Characteristics and Fidelity: Enabling Skills, Interests, Aptitudes and Abilities</u>

Among the many factors which can influence the success and effectiveness of training, few are as important as the aptitudes, abilities, attitudes, competencies and motivational structures that an individual brings into the training situation. Indeed, how a trainee mediates the acquisition, organization and retention of knowledges and skills is often a direct function of such state variables. From the training systems designer's perspective such facts present at times what appears to be a

hopeless tangle, particularly when he must design training programs for large numbers of trainees, differing in personal characteristics and abilities. Achieving a balance between individual differences, training aids/devices, conditions and curricula is not an easy task. Nevertheless, much more could be done to assist and inform those charged with making sure that methods, media and devices are appropriate ones and effective with a wide range of aptitudes, abilities, intelligences, etc. With this in mind, a continuing aim should be to examine the relationships between individual differences variables, problem-solving strategies, simulator fidelity levels and transfer performance. In our own work, a general strategy has been to collect individual difference measures from all subjects. While we believe this should continue, studies might also be designed in terms of general aptitude-training-interaction models (ATI). Such models represent a longstanding conceptual approach for research aimed at fitting instructional treatments to individuals (Cronbach & Snow, 1977).

B.2 Trainee Strategies

While all training situations require individuals to adopt certain general strategies appropriate for learning (maintain the "right" attitude, "be motivated", "pay attention", etc.), the learning of particular skills often requires that other, more task-specific strategies also be used. Such is the case with troubleshooting. With respect to electronics troubleshooting, for example, a number of specific strategies have been identified over the years. A few include: reliability strategies, conditional probability strategies, syndrome analysis, signal

tracing, split-half approach, etc. (Van Cott & Kinkade, 1972).

Apart from the task itself, there are many other factors which may influence strategy selection, including such things as trainee intelligence, aptitudes, abilities, previous experience, instructions, etc. One of our aims has been to not only identify the types of strategies used by subject, but to relate them to individual difference variables and specific training experiences. We have noted, for example, that three general strategies are used by trainees to solve problems during transfer, viz. a logical strategy, a sequential strategy and a haphazard strategy. However, it should be noted that so far only the most preliminary examination of these strategies has been done. How they covary with other variables within the area of training device fidelity has yet to be determined. Certainly more research is needed to examine this aspect of trainee behavior in more detail so that answers to questions like the following might be obtained. Are some simulator training conditions more apt to result in the use of more efficient problem-solving strategies than others? What is the relationship between trainee aptitudes, abilities, etc. and strategy selection? Can strategies be pretrained? How are trainee abilities/aptitudes, strategies and transfer performance related? Do problem-solving strategies change during the transfer process? And if so, how?

B.3 Generic Troubleshooting Training

An important question which bears on the overall strategy used to teach common skills across a variety of tasks and jobs, concerns the effectiveness of so-called "generic", as opposed to "task-specific" training. There are, a number of potential

advantages to such an approach, not the least of which are: reduced costs, reduction in the number of training programs and curricula, the need for fewer instructors, less duplication of equipment, and so on. With respect to the training of troubleshooting behavior, an especially good case can be made (at least at the intuitive level) for incorporating a generic approach, particularly since it is a skill which cuts across a wide variety of maintenance activities and tasks.

The study of generic (sometimes called, "context-free") training is not new of course, however most investigations have examined transfer from computer-based situations to "real-world" tasks (Rasmussen & Rouse, 1981). Few have looked at transfer between different training devices which, though sometimes dramatically dissimilar in terms of appearance and function, share an emphasis on certain skills and behaviors. In the area of maintenance trainers, for example, such research might help identify skills which were non-device specific, and in so doing, suggest more efficient means of allocating training resources.

C. Training Device Design Issues

C.1 Amount of Training and Simulator Fidelity

An essential condition for the learning of any skill is practice. This is no less true for skills which might be taught by training devices or training simulators. Indeed, one of the principle advantages of training simulators is that they often permit more opportunities for practice when compared with operational equipment. For example, with many simulators, critical skills can be repeatedly practiced until a desired level of performance is obtained. However, as essential as practice

is, it is still not possible to specify exactly how much practice is needed for maximum transfer of training. To be sure, we know that, as a general rule, more practice results in more transfer, but beyond this we can say little more. Moreover, our knowledge is even less when simulator fidelity levels are added to the equation. For example, is less practice needed when high levels of training device fidelity are used? And how does amount of practice on such a device interact with other important training sytems variables, such as trainee ability, task complexity, task type, etc? Do some trainees benefit more from increased practice than others, regardless of simulator fidelity level? It is obvious that answers to questions like these are critical for appropriate training device design.

C.2 Costs of Maintenance Actions and Decisions

In the real world of maintenance and repair, costs, in terms of money, time or both, are almost always associated with technician actions and decisions. This is no less true in the military services where, in addition to the monetary and temporal costs, it is also reflected by such indices as effectiveness and combat readiness. There are many areas where such costs might be trimmed, however, one area which has received only passing consideration is the teaching of trainee "awareness" of costs during training. Indeed, instructional programs which incorporate simulator-based training would appear to offer unusual opportunities in this respect, for unlike other modes of training, in which an awareness of what actions and decisions cost is often reduced to an abstraction, the very nature of working with devices which resemble "the real thing" makes the idea of teaching an

awareness of cost a realistic goal. Studies might focus on such things as: the general effects of costing upon trainee strategy selection and performance; the differential effects of fidelity and cost restrictions upon learning transfer; a comparison of two methods of teaching "cost awareness", etc.

C.3 Set and Pretraining

It is an old and well-established fact that the "set" of a learner can often significantly affect the outcome of training (Harlow, 1949; Wolfe, 1951). That is, how an individual enters a learning situation may have positive or negative consequences on his learning performance depending on such things as: attitude, previous experience with similar situations, expectations about the situation in general, and so on.

A particularly effective means of altering a trainee's "set" is via instructions given prior to training. Such a procedure may be viewed as a form of "pretraining" in which an individual's way of thinking and responding to the training situation is deliberately changed to more closely meet some experimental criterion established by the trainer. Used in an experimental setting, pretraining can be an effective means of isolating important variables and their interactions. For example, in the present context of simulator fidelity, instructional set (pretraining), when manipulated along with other factors, might be used to help identify nonspecific aspects of transfer which might serve the development of selection criteria for matching trainees to simulator training conditions.

D. Training Device Implementation Issues

D.1 User Acceptance of a Training Device

It is perhaps a truism, but even the most well-designed and appropriate training device can not be expected to provide adequate, much less effective, training if it is not used properly. Although on can think of many reasons why a device might be under- or improperly utilized (lack of knowledge by the trainer, poor manuals/documentation, ect.), nonacceptance of a training device by instructors and/or trainees has been shown to be a frequent cause (Semple, 1981; Caro, Shelnutt & Spears, 1981; Bleda, 1979). Despite this, however, very few studies have addressed the issue directly (Biersner, 1975; Stoffer, Blaiwes, 1980). As a result, we know little about such things as: the relationship between simulator fidelity level and user acceptance; the effects of instructor attitude toward a training device on trainee motivation and performance; how training device realism motivates training during learning; how misperceptions on the part of trainers/trainees affect transfer performance, etc. Clearly, more research is needed.

E. Training Effectiveness Assessment Issues

E.1 Retention of Learning Following Training

An important criterion for assessing training effectiveness is the extent to which learning obtained during training is retained by personnel once instruction is complete. This is especially true when costs of training are high and when maintenance of adequate performance skills over varying intervals of time is critical. One aspect of the general retention issue which is particularly pertinent to simulator- and other device-

based training programs is the extent to which delays between training device and system experiences influence transfer performance. One might guess that zero-delay would be best, with longer intervals producing poorer and poorer transfer. However, in the real world, it is not always possible to give trainees immediate experience with operational systems, and so delays between simulator training and actual equipment training are more often the rule than an option. For this reason, the issue of delay remains an important and proper question for study.

A second aspect of retention which is germane not only to device-based training programs, but to training curricula at-large, is the degree to which skills are maintained after training of <u>all</u> kinds have been completed. With respect to device-based training, this frequently refers to the period following both simulator and actual equipment (OJT) training. A number of questions come to mind. Are some simulator/AET experiences more likely to produce better long-term retention of skills than others? If so, what particular features of those experiences are most important for long-term skills maintenance? What is the proper mix of simulator and AET training for ensuring adequate long term skills maintenance? In terms of long-term retention, what advantages do certain simulator variations have over others?

F. <u>Expected Products and Outcomes</u>

In summary, we believe that a number of important general and specific products might be expected from a research program like the one described.

First, the results of such work would help further clarify the general relationship between simulator fidelity and training efficiency, and even more importantly, delineate this relationship in terms of a number of new task domains, task and critical training systems variables. Second, we feel that the findings would significantly broaden the empirical data base upon which new guidelines and principles might be derived. Such guidelines and principles, in turn, would assist those involved in the design, development and implementation of simulator-based training systems of every sort. Third, answers to many important training systems issues would result. A few representative ones include:

1. To what extent are aptitude-treatment-interaction models applicable to simulator-based training?
   -- Can particular modes of instruction (simulator training conditions) be matched to individual differences variables so that persons are assigned to the most appropriate training procedure?
2. How do certain specific instructional strategies, individual differences and simulator fidelity levels interact?
   -- Is the performance of subjects who learn a "cookbook" approach to problem solving any less reliable than subjects who are taught only operational principles?
   -- Are "less analytic" trainees more likely to benefit from a "cookbook" approach to troubleshooting than their "more analytic" counterparts?

-- Do some simulator training conditions result in the use of more efficient troubleshooting strategies than others?

-- Can troubleshooting strategies be pretrained?

3. To what extent do troubleshooting skills learned in one maintenance training situation transfer to another maintenance training situation?

   -- How is simulator fidelity level related to transfer across training tasks sharing common troubleshooting skills?

4. How does the amount of training one receives using a simulator interact with individual differences and simulator fidelity levels to affect such things as transfer performance, retention, etc?

   -- Is less practice needed when high levels of simulator fidelity are used?

   -- Are some trainees more likely to benefit more from increased simulator training than others?

   -- What are the long-term effects of increased simulator training?

5. What are the effects of "costing" trainees for their actions and decisions during the training?

   -- Are some "costing" methods more effective than others with respect to engendering the use of cost-effective strategies among trainees?

   -- How does increased "cost awareness" on the part of trainees related to transfer performance?

   -- Can "cost awareness" be taught?

6. How do instructor attitudes toward simulators and simulator-based training affect trainee learning and transfer performance?
   -- Do "negative" attitudes on the part of instructors seriously impair trainee learning and transfer performance?
7. How do delays between simulator training and transfer tests on operational equipment affect retention of skills?
   -- Are some simulator fidelity levels associated with better retention of skills than others?

Such a list, of course, only begins to sample the questions and issues which need answers. Nevertheless, we believe that research designed to address some of them would go a long way toward the goal of providing useful empirically-based guidelines to those charged with making sure that training programs and devices are as cost-effective and efficient as possible.

References

Adams, J. A. (1979). On the evaluation of training devices. Human Factors, 21, (6), 711-720.

Ayres, A., Hays, R. T., Singer, M. J. & Heinicke, M. (1984). An annotated bibliography of abstracts on the use of simulators in technical training. (ARI Research Product 84-21). Alexandria, VA: U.S. Army Research Institute. (AD A156 792)

Baum, D. R., Riedel, S., Hays, R. T. & Mirabella, A. (1982). Training effectiveness as a function of training device fidelity (ARI Technical Report 593). Alexandria, VA: U.S. Army Research Institute. (AD A140 997)

Biersner, R. J. Attitudes and other factors related to aviation maintenance training effectiveness (CNETS Rep. 6-75). Pensacola, FL: Chief of Naval Education and Training Support. December 1975.

Bilodeau, I. M. (1966). Informational feedback. In E. A. Bilodeau (Ed.), Acquisition of Skill. New York: Academic Press.

Bleda, P. R. Realtrain improves soldier attitudes toward the Army (ARI Tech. Paper 363). Alexandria, VA: U. S. Army Research Institute, May 1979. (AD A072 334)

Bresnard, G. G. & Briggs, L. J. (1956). Comparison of performance upon the E-4 fire control systems simulator and upon operational equipment. (Development Report: AFPTRC-TN56-47). Lowery Air Force Base, CO: Armament Systems Personnel Research Laboratory, AFPTRC.

Bunker, W. M. (1978). Training effectiveness versus simulation realism. In the Eleventh NTEX/Industry Conference Proceedings. Naval Training Equipment Center: Orlando, FL.

Caro, P. W., Shelnutt, J. B., & Spears, W. D. Aircrew training devices: Utilization (Tech. Rep. AFHRL-TR-80-35). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, January 1981.

Cicchivelli, L. F., Harmon, K. R., Keller, R. A., & Kottenstette, J. P. (1980). Relative cost and training effectiveness of the 6883 three-dimensional simulator and actual equipment. Air Force Human Resources Laboratory Report No. AFHRL-TR-80-24.

Cronbach, L. J. & Snow, R. E. (1977). Aptitudes and Instructional Methods. Irvington: New York

Ellis, H. C. (1965). Transfer of Learning. The Macmillan Company: New York.

Finch, C. R. (1971). Troubleshooting instruction in vocational technical education via dynamic simulation. University Park, PA: Pennsylvania State University, (Eric Document Reproduction Service No. ED 055 204).

Ghiselli, E. R., Campbell, J. P., & Zedeck, S. (1981). Measurement Theory for the Behavioral sciences. San Francisco: Freemen.

Glass, A. A. (1967). Problem solving techniques and troubleshooting simulations in training electronic repairmen. (Doctoral dissertation, Columbia University). Dissertation Abstracts, 28, 1717B.

Harlow, H. F. (1949) The formation of learning sets. Psychological Review, 56, 51-65.

Hays, R. T. (1980). Simulation Fidelity: A Concept Paper. Technical Report 490. U. S. Army Research Institute for the Behavioral and Social Sciences. (AD A101 475)

Hays, R. T. & Singer, M. J. (1983) Research Issues in Training Device Design: The Organization of a Data Base. Technical Report 588. U. S. Army Research Institute for the Behavioral and Social Sciences. (AD A140 815)

Henneman, A. & Rouse, W. (1984). Measures of human problem-solving performance in fault diagnosis tasks. IEEE Transaction on Systems, Man, and Cybernetics. Vol. SMC-14, 99-112.

Johnson, W. G. & Rouse, W. B. (1982). Training maintenance technicals for troubleshooting: Two experiments with computer simulations. Human Factors, 24(3), 271-276.

Johnson, W. B. & Fath, J. L. (1983). Implementation of a mixed-fidelity approach to maintenance training. (First year report Contract No. MDA 903-82-CC-0354). Norcross, GA: Search Technology, Inc.

Kinkade, R. G. & Wheaton, G. R. (1972). "Training Device Design", in H. P. Van Cort & R. G. Kinkade (Eds.), Human Engineering Guide to Equipment Design, U. S. Government Printing Office, Washington, DC 20402.

Mirabella, A. & Wheaton, G. (1974). Effects of task index variations on transfer of training criteria (Tech. Rep. NAVTRAEQUIPCEN-72-C-0126-1). Orlando, FL: Navy Training Equipment Center.

Provenmire, K. & Roscoe, S. (1971). An evaluation of ground-based flight trainers in routine primary flight training. Human Factors, 13(2), 109-116.

Rasmussen, J. & Rouse, W. (1981). Human Detection and Diagnosis of System Failures. New York: Plenum.

Rigney, J. W., Towne, D. M. King, C. A. & Moran, P. J. (1978). Field evaluation of the generalized maintenance trainer-simulator: I--Fleet communications system (Tech. Rep. No. 89). Arlington, VA: Office of Naval Research.

Rouse, S. & Rouse, W. (1982). Cognitive style as a correlate of human problem solving in fault diagnosis tasks. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-12, 649-652.

Semple, C. A., Hennessy, R. T., Sanders, M. S., Cross, B. K., & McCauley, M. E. Aircrew training devices: Fidelity features (Tech. Rep. AFHRL-TR-80-36). Brooks Air Force Base, TX: Air Force Human Resources Laboratory, January 1981.

Stoffer, G. R., Blaiwes, A. S., & Brictson, C. A. User acceptance of R&D in Navy training: The problem, major constraints and an initial model of the acceptance process. Proceedings of the First Interservice/Industry Training Equipment Conference. Orlando, FL November 1980, 17-30.

Unger, K. W., Swezey, R. W., Hays, R. T. & Mirabella, A. (1984). Army Maintenance Training and Evaluation Simulation System (AMTESS) device evaluation: Volume II, transfer-of-training assessment of two prototype devices (ARI Technical Report 643). Alexandria, VA: U.S. Army Research Institute. (AD A159 516)

VanCott, H. P. & Kinkade, R. G. (1972). Human Engineering Guide To Equipment Design. U. S. Government Printing Office, Washington, DC 20402.

Wheaton, G. R. & Mirabella, A. (1972). Effects of task index variations on training effectiveness criteria. Technical Report: NAVTRAEQUIPCEN 71-C-0059-1. Orlando, FL: Navy

Wolfe, D. Training. In S. S. Stevens (Ed.), Handbook of Experimental Psychology, Wiley, 1951, 1267-1286

END